

## Supplementary Materials

**Title: Genome-wide association study meta-analysis of neurofilament light (NfL) levels in blood reveals novel loci related to neurodegeneration**

### Author's List:

Shahzad Ahmad\*<sup>1,2</sup>, Mohammad Aslam Imtiaz\*<sup>3</sup>, Aniket Mishra<sup>6</sup>, Ruiqi Wang<sup>7</sup>, Marisol Herrera-Rivero<sup>8,9</sup>, Joshua C Bis<sup>10</sup>, Myriam Fornage<sup>11</sup>, Gennady Roshchupkin<sup>1</sup>, Edith Hofer<sup>12,13</sup>, Mark Logue<sup>14,15</sup>, WT Longstreth Jr<sup>16</sup>, Rui Xia<sup>11</sup>, Vincent Bouteloup<sup>6</sup>, Thomas Mosley<sup>17</sup>, Lenore J Launer<sup>18</sup>, Michael Khalil<sup>19</sup>, Jens Kuhle<sup>20</sup>, Robert A. Rissman<sup>21</sup>, GenevieveChene<sup>6</sup>, Carole Dufouil<sup>6</sup>, Luc Djoussé<sup>22</sup>, Michael J. Lyons<sup>23</sup>, Kenneth J. Mukamal<sup>24</sup>, William S. Kremen<sup>25</sup>, Carol E. Franz<sup>25</sup>, Reinhold Schmidt<sup>12</sup>, Stephanie Debette<sup>6, 26</sup>, Monique M.B.Breteler<sup>3,5</sup>, Klaus Berger<sup>27</sup>, Qiong Yang<sup>7</sup>, Sudha Seshadri<sup>28,7</sup>, N. Ahmad Aziz<sup>3,4, #</sup>, Mohsen Ghanbari<sup>1, #</sup>, M. Arfan Ikram<sup>1, #</sup>

### Author's Affiliations:

- <sup>1</sup> Erasmus University Medical Center, PO Box 2040, 3000 CA Rotterdam, the Netherlands.
- <sup>2</sup> Oxford-GSK Institute of Computational and Molecular Medicine (IMCM), Centre for Human Genetics, Nuffield Department of Medicine (NDM), University of Oxford, Oxford, OX3 7BN, UK
- <sup>3</sup> Population Health Sciences, German Center for Neurodegenerative Diseases (DZNE), Venusberg-Campus 1/99, 53127, Bonn, Germany.
- <sup>4</sup> Department of Neurology, Faculty of Medicine, University of Bonn, 53127, Bonn, Germany.
- <sup>5</sup> Institute for Medical Biometry, Informatics and Epidemiology (IMBIE), Faculty of Medicine, University of Bonn, Venusberg-Campus 1, 53127 Bonn, Germany.
- <sup>6</sup> University of Bordeaux, Inserm, Bordeaux Population Health Research Center, UMR 1219, F-33000 Bordeaux, France
- <sup>7</sup> Boston University, Boston, MA, 02215, United States.
- <sup>8</sup> Department of Genetic Epidemiology, Institute of Human Genetics, University of Münster, Münster, Germany.
- <sup>9</sup> Department of Psychiatry, University of Münster, Münster, Germany.
- <sup>10</sup> Cardiovascular Health Research Unit, Department of Medicine, University of Washington, 1730 Minor Ave #1360, Seattle, WA 98101, USA.
- <sup>11</sup> Brown Foundation Institute of Molecular Medicine, McGovern Medical School, University of Texas Health Science Center at Houston, 1825 Pressler Street Houston, 77030 TX, USA.

- <sup>12</sup> Clinical Division of Neurogeriatrics, Department of Neurology, Medical University of Graz, Auenbruggerplatz 22, 8036, Graz, Austria.
- <sup>13</sup> Institute for Medical Informatics, Statistics and Documentation, Medical University of Graz, Auenbruggerplatz 2, Fifth Floor, 8036, Austria.
- <sup>14</sup> National Center for PTSD, Behavioral Sciences Division at VA Boston Healthcare System, Boston, 150 South Huntington Avenue Boston, MA 02130, USA.
- <sup>15</sup> Department of Psychiatry and Biomedical Genetics, Boston University School of Medicine, Boston, 72 East Concord Street E200 Boston, MA 02118, USA.
- <sup>16</sup> Departments of Neurology and Epidemiology, University of Washington, Seattle, 3980 15th Ave NE Seattle, WA 98195, USA.
- <sup>17</sup> MIND Center, University of Mississippi Medical Center, Jackson, 2500 North State Street Jackson, MS 39216, USA.
- <sup>18</sup> Laboratory of Epidemiology and Population Science, NIA Intramural Research Program, 251 Bayview Blvd, Baltimore MD, 21224, USA.
- <sup>19</sup> Department of Neurology, Medical University of Graz, Auenbruggerplatz 22, 8036, Graz, Austria.
- <sup>20</sup> Research Center for Clinical Neuroimmunology and Neuroscience University Hospital, Spitalstrasse 2, CH-4031 Basel Switzerland .
- <sup>21</sup> Department of Physiology and Neuroscience, Alzheimer's Therapeutic Research Institute, Keck School of Medicine of the University of Southern California.
- <sup>22</sup> Brigham and Women's Hospital, Harvard Medical School, Boston, 75 FRANCIS STREET, BOSTON MA 02115, MA, USA.
- <sup>23</sup> Department of Psychological & Brain Sciences, Boston University, Boston, 64 Cummington Mall # 149, Boston, MA 02215, USA.
- <sup>24</sup> Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, 330 Brookline Avenue Boston, MA 02215, USA.
- <sup>25</sup> Department of Psychiatry and Center for Behavior Genetics of Aging, University of California, San Diego, La Jolla, CA, 92093, USA.
- <sup>26</sup> CHU de Bordeaux, Department of Neurology, Institute for Neurodegenerative Diseases, F-33000 Bordeaux, France.
- <sup>27</sup> Institute of Epidemiology and Social Medicine, University of Münster, Münster, Institut für Epidemiologie und Sozialmedizin Albert-Schweitzer-Campus 1, Gebäude D3 48149, Germany.
- <sup>28</sup> Glenn Biggs Institute for Alzheimer's and Neurodegenerative Diseases, University of Texas Health Sciences Center, San Antonio, Texas, USA.

## **Supplementary Notes:**

### **Participating Cohort Description**

#### **The Alzheimer's Disease Neuroimaging Initiative (ADNI) Study**

The ADNI was launched in 2003 as a public-private partnership, led by Principal Investigator Michael W. Weiner, MD. ADNI is a global research study that actively supports the investigation and development of treatments that slow or stop the progression of Alzheimer's disease. In this multisite longitudinal study, researchers at 63 sites in the US and Canada track the progression of AD in the human brain with clinical, imaging, genetic, and biospecimen biomarkers through the process of normal aging, early mild cognitive impairment (EMCI), and late mild cognitive impairment (LMCI) to dementia or AD. The overall goal of ADNI is to validate biomarkers for use in Alzheimer's disease clinical treatment trials. The primary goal of ADNI has been to test whether serial magnetic resonance imaging (MRI), positron emission tomography (PET), other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of mild cognitive impairment (MCI) and early Alzheimer's disease (AD). For up-to-date information, see [www.adni-info.org](http://www.adni-info.org).

Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database ([adni.loni.usc.edu](http://adni.loni.usc.edu)). Phenotypic and genetic data were downloaded from the data repository hosted at the Laboratory of Neuroimaging (LONI) at the University of Southern California, the LONI Image & Data Archive (IDA). Principal Component Analyses (PCA) were performed using Eigensoft based on pruned genetic data, with exclusion of complex and HLA regions. Ethnic outliers were removed based on 6SD from the mean. Plasma tau was analyzed with the Human Total Tau kit (research use only grade, Quanterix, Lexington, MA) on the Simoa HD-1 analyzer (CE marker). ADNI1 SNP genotype data were used to perform the GWAS (Illumina Human610-Quad BeadChip). Only non-Hispanic whites ADNI1 participants were included in the GWAS of circulating tau levels. Winsorizing at 4 SD was used to removed outliers. QC on genetic data was performed (call-rate > 0.99, P-value of HWE > 10<sup>-4</sup>; MAF > 1%). Calculation of an empirical kinship matrix was performed to account for

relatedness in the association analyses. Linear mixed-effects models were used to evaluate the association of genetic variants with circulating total-tau levels, adjusted for age, sex, and PCs.

### **The Rotterdam Study (RS1 and RS2)**

The Rotterdam Study is a population-based cohort study among inhabitants of a district of Rotterdam (Ommoord), the Netherlands, that aims to examine the determinants of disease and health in the elderly with a focus on neurogeriatric, cardiovascular, bone, and eye disease.<sup>8</sup> In 1990-1993, 7,983 persons aged  $\geq 55$  years participated and were re-examined every 3 to 4 years (Rotterdam Study I). In 2000-2001, the cohort was expanded by 3,011 persons who were of the same age but had not yet been part of the Rotterdam Study (Rotterdam Study II) and recently moved into the area. All participants had blood collected during their first center visit, which was followed by DNA extraction.

**Genotyping and imputations:** Genotyping was done in participants with high-quality extracted DNA in 2007-2008 and was performed at the Human Genotyping Facility, Genetic Laboratory Department of Internal Medicine, Erasmus MC, Rotterdam, The Netherlands. Imputation of SNPs was established using the Michigan Imputation server and the HRC reference panel. More specifically, the SHAPEIT2 software was used (v2.r790) to phase the data and Minimac 3 was employed for imputation to the HRC reference panel (v1.0). QC included deletion of participants with a genotype completion rate ( $<90\%$ ), a low genotype call rate ( $<95\%$ ), sex-mismatches, duplicate pairs (just one participant), uncalled variants in over 5% of the individuals and significant violations of the expected Hardy–Weinberg Equilibrium proportions ( $P < 10^{-6}$ ).

**A $\beta$ , total-tau and NfL Assessment of plasma:** Blood was sampled in EDTA treated containers and centrifuged. Subsequently plasma was aliquoted and frozen at  $-80^{\circ}\text{C}$  according to standard procedures. Measurements were done in two separate batches. The first batch included in total 2,000 samples, obtained from a random selection of 1,000 participants from sub-cohort RS-I-4 and 1,000 from RS-II-2. The second batch included in total 3,094 samples from the remaining participants. All measurements were performed at Quanterix (Lexington, MA, USA) on a single molecule array (Simoa) HD-1 analyzer platform.<sup>1</sup> Samples were tested in duplicate. Two quality control (QC) samples were run on each plate for each analyte. NfL was measured with the NF-light advantage kit.<sup>2</sup> The Simoa Human Neurology 3-Plex A assay (N3PA) was used for

measuring the concentration of total-tau, A $\beta$ -42 and A $\beta$ -40. When duplicates or single measurements were missing or in the case the concentration coefficient of variation (CV) exceeded 20% or control samples were out of range, participant's data were excluded from the analyses<sup>3</sup>.

### **The Rhineland Study**

The Rhineland Study is an ongoing single-center, population-based prospective cohort study among people aged 30 years and above in Bonn, Germany. All individuals living in two pre-defined recruitment areas are invited to participate in the study, and they are predominantly German of Caucasian descent. The only exclusion criterion is an insufficient command of the German language to provide informed consent. A primary objective of the Rhineland Study is to identify determinants and markers of healthy and unhealthy aging. The study uses a deep-phenotyping approach and has a special focus on brain aging. At baseline, participants complete an 8-hour in-depth multi-domain phenotypic assessment and various types of biomaterials (blood, urine, stool, and hair) are collected. Approval to undertake the study was obtained from the ethics committee of the University of Bonn, Medical Faculty. We obtained written informed consent from all participants in accordance with the Declaration of Helsinki.

***NFL quantification:*** Blood samples were collected in Vacutainer EDTA tubes and centrifuged at 2000 x g for 10 min at room temperature. The obtained plasma samples were aliquoted and stored at -80 °C. NfL levels were assessed using the Simoa<sup>®</sup> NF-light Kit (103186) in a HD-1 Analyzer (Quanterix, Billerica, USA), following manufacturer's instructions. Briefly, plasma samples were thawed with the help of the 96 FluidX Tube Rack Thawing Station (Brooks Life Sciences, Darmstadt, GE) and centrifuged at 3000 x g for 10 min at 4 °C. On-board the instrument, prepared samples were diluted using the standard 4x dilution protocol and incubated with target antibody coated paramagnetic beads and biotinylated detector antibody, followed by addition of streptavidin- $\beta$ -galactosidase enzyme complex. If NfL has been captured and labeled on the bead, the enzyme hydrolyzes the resorufin  $\beta$ -D-galactopyranoside substrate into a fluorescent product, which provides the signal for

measurement. The NfL concentrations in pg/mL were assessed from the calibration curve generated for each assay. Intra and inter-assay coefficients of variation for plasma controls were lower than 20%, in accordance with manufacturer's recommendation.

**Genotyping:** DNA extracted from buffy coat samples were genotyped using Infinium Omni2.5 Exome-8 BeadChip containing 2,612,357 SNPs and processed using GenomeStudio (version 2.0.5). Quality control of genotypes was performed using PLINK software (version 1.9). We used 1000 Genomes phase 3 reference panel using hg19 version for imputation of missing genotypes using impute2 software. GWAS association methods was performed using rvtest software and using standard score test.

### **The Cardiovascular Health Study (CHS) cohort**

The Cardiovascular Health Study (CHS) is a population-based cohort study of risk factors for coronary heart disease and stroke in adults  $\geq 65$  years conducted across four field centers [PMID: 1669507]. The original predominantly European ancestry cohort of 5,201 persons was recruited in 1989-1990 from random samples of the Medicare eligibility lists; subsequently, an additional predominantly African-American cohort of 687 persons was enrolled for a total sample of 5,888. Blood samples were drawn from all participants at their baseline examination and DNA was subsequently extracted from available samples. CHS was approved by institutional review committees at each field center and individuals in the present analysis had available DNA and gave informed consent including consent to use of genetic information for the study of cardiovascular disease. For this ancillary study, all participants who underwent routine oral glucose tolerance testing at the 1996-1997 clinic visit were included. Entry criteria for the OGTT included in-person attendance in 1996-1997, fasting status, and absence of anti-diabetic medication.

**NFL quantification:** Serum NfL was measured in singlet with the Quanterix single molecule array platform at the CHS Central Laboratory at the University of Vermont using the HD-X analyzer and the Simoa Human Neurology 4-Plex A assay. Preliminary studies on ~200 duplicate samples

demonstrated very high reproducibility. The detectable range was 0.416 - 1648 pg/mL. Inter-assay coefficients of variation ranged from 8.53 – 10.68%.

**Genotyping:** Genotyping was performed at the General Clinical Research Center's Phenotyping/Genotyping Laboratory at Cedars-Sinai among CHS participants who consented to genetic testing and had DNA available using the Illumina 370CNV BeadChip system (for European ancestry participants, in 2007) or the Illumina HumanOmni1-Quad\_v1 BeadChip system (for African-American participants, in 2010). European ancestry participants were excluded from the GWAS study sample due to the presence at study baseline of coronary heart disease, congestive heart failure, peripheral vascular disease, valvular heart disease, stroke or transient ischemic attack or lack of available DNA. Beyond laboratory genotyping failures, participants were excluded if they had a call rate  $\leq 95\%$  or if their genotype was discordant with known sex or prior genotyping (to identify possible sample swaps). After quality control, genotyping was successful for 3,268 European ancestry and 823 African-American participants.

**SNP exclusions** In CHS, the following exclusions were applied to identify a final set of 306,655 autosomal SNPs: call rate  $< 97\%$ , HWE  $P < 10^{-5}$ ,  $> 2$  duplicate errors or Mendelian inconsistencies (for reference CEPH trios), heterozygote frequency = 0, SNP not found in HapMap. Imputation to the HRC r1.1 2016 panel was performed on the Michigan imputation server. SNPs were excluded for variance on the allele dosage  $\leq 0.01$ .

### **The BiDirect cohort**

The BiDirect Study was initiated in 2009 as a prospective, observational study integrating three cohorts: 1) community-dwelling adults (control diagnostic group), 2) patients with an acute depressive episode (depression diagnostic group), and 3) patients who recently suffered from acute myocardial infarction (MI diagnostic group). The study, whose principal goal is the exploration of the bidirectional relationship between depression and subclinical arteriosclerosis, recruited participants in the district of Münster, Germany, and carried out extensive phenotyping and follow-up of all cohorts in parallel. The study design and methods have been described in detail in Teismann et al., 2014<sup>4</sup>. All participants

provided informed written consent. The BiDirect Study was approved by the ethics committee of the University of Münster and the Westphalian Chamber of Physicians in Münster, North-Rhine-Westphalia, Germany, and is funded by the German Federal Ministry of Education and Research (grant numbers FKZ-01ER0816 and -01ER1506).

**NfL quantification:** Quantification of serum NfL was conducted at the University Hospital Basel, Switzerland, using the single molecule array (Simoa®) HDX analyzer (Quanterix, Lexington, MA, USA). Measurements of serum NfL were obtained from non-fasting blood samples collected at the first visit, using the Simoa® NF-light Advantage Kit. The sNfL values obtained at initial assessment were log<sub>2</sub>-transformed and used for analysis.

**Genotyping:** Genomic DNA was isolated from whole blood samples with EDTA using standard DNA extraction kits and procedures at the University of Münster. Whole-genome genotyping was performed with the Infinium PsychArray BeadChip v1 (Illumina) at Life&Brain GmbH (Bonn, Germany). Following basic quality control of the genotyped data (HWE  $p > 10^{-6}$ , MAF  $\geq 1\%$ , SNP and sample call rate  $\geq 95\%$ ), imputation was performed with SHAPEIT and IMPUTE2 using the 1000 Genomes Project, phase 3, European population reference panel. Individuals were removed from the sample based on relatedness, outlier detection and missing serum NfL data. Genomic coordinates are given in the hg19 genome build. Association analysis method: An additive linear regression model adjusting for Age, Sex, Diagnostic group (depression/MI/control) and first 20 Principal Components was applied using Plink 1.9.

### **The Framingham Heart Study (FHS)**

The FHS is a single-site, community-based, ongoing cohort study that was initiated in 1948 to investigate prospectively the risk factors for CVD. The Original cohort (n=5209) were randomly recruited in the town of Framingham, MA, USA, and examined every 2 year since 1948; their children and spouses of the children (n=5124), the Offspring cohort, were enrolled in 1971 and examined approximately every 4 year since 1971. This study includes 2048 Offspring subjects. Written informed consent to genetic research has been obtained on all individuals included



in this study. Ethics permission for FHS and genetic research in FHS was obtained from the Institutional Review Board of Boston University Medical Campus (IRB number H-32132, H-26671).

**NFL quantification:** NFL levels were quantified using the the Quanterix 4 plex and assays at University of Vermont by Russ Tracy.

**Genotyping:** Genotyping was performed on Affymetrix 500K mapping array plus Affymetrix 50K supplemental array. Pre-imputation QC filter include excluding genotyped SNPs call rate<97%, HWE  $P < 10^{-6}$ , MAF<1%, Number of Mendelian errors $\geq$ 1000, or at locations that did not map to GRCh37. Genotyping data was imputed using Haplotype Reference Consortium panel (hg19) using university of Michigan imputation server. Association analysis was performed using R software, lme4 function.

### **The Austrian stroke prevention family study (ASPS-Fam)**

The Austrian stroke prevention family study (ASPS-Fam) is a prospective single-center, community-based study on the cerebral effects of vascular risk factors in the normal elderly population of the city of Graz, Austria. The ASPS-Fam represents an extension of the Austrian stroke prevention study (ASPS), which was established in 1991<sup>5,6</sup>. Between 2006 and 2013, study participants of the ASPS and their first grade relatives were invited to enter ASPS-Fam. Inclusion criteria were no history of previous stroke or dementia and a normal neurologic examination. The entire cohort underwent an extended diagnostic work-up including clinical history, blood tests, cognitive testing, MR imaging and a thorough vascular risk factor assessment. A total of 419 individuals were included into the study and 287 participants from 153 families had GWAS and NFL data. The study protocol was approved by the ethics committee of the Medical University of Graz, Austria, and written informed consent was obtained from all subjects<sup>7,8</sup>.

**NFL quantifications:** All serum samples were analyzed at the University Hospital Basel, Switzerland. Serum NfL levels were determined by single molecule array (Simoa) assay using the capture monoclonal antibody (mAB) 47:3 (initial dilution 0.3 mg/mL; Art. No. 27016) and the biotinylated detector mAB 2:1 (0.1  $\mu$ g/mL; Art. No. 27018) from Uman Diagnostic<sup>9</sup> transferred onto the Simoa platform<sup>10</sup>. The samples from the

same participants were analyzed together in the same run to avoid within-subjects run-to-run variability. Intra- and interassay variability of the measurements were evaluated with three native serum samples in five consecutive runs on independent days. The mean coefficients of variation (CVs) of duplicate determinations for concentration were 8.5% (9.5 pg mL<sup>-1</sup>, sample 1), 5.4% (23.2 pg mL<sup>-1</sup>, sample 2) and 7.8% (98.5 pg mL<sup>-1</sup>, sample 3). Interassay CVs for serum were 7.8% (sample 1), 8.3% (sample 2) and 4.9% (sample 3)<sup>11</sup>.

**Genotyping:** Genotyping was performed using the Affymetrix Genome-Wide Human SNP Array 6.0. Genotypes were imputed using HRC (r1.1 2016) imputation panel using Michigan Imputation Server. Pre-imputation quality control include sample call rate 98%, SNP call rate 98%, MAF 0.05, HWE p-value < 5\*10<sup>-6</sup>. Other sample exclusion criteria include sample failures, sex mismatch, cryptic relatedness, heterozygosity: mean±3SD.

#### **The MEMENTO cohort**

The MEMENTO study prospectively included 2,323 individuals of French memory clinic, who presented either isolated subjective cognitive complaints (SCCs) or mild cognitive impairment (MCI; defined as test performance 1.5 SD below age, sex and education-level norms) while not demented (Clinical Dementia Rating [CDR] <1), between January 2011 to June 2014. Participants were followed every 6 months for 5 years<sup>12</sup>.

**NFL profiling:** NFL profiling was carried out using Simoa™ NF-light kit on a Quanterix H1 analyser (Quanterix, MA, USA). Intra CV: 12.7%

**Genotyping and imputation:** Pre-imputation QC includes removal of SNPs with MAF<0.01, callrate<0.98 and HWE<0.001; Removed samples with callrate<0.05, heterozygosity beyond 3SD, failed sex-check using genotype data of X-chromosome, relates sample based on IBD ( $\pi_{\text{hat}} > 0.1875$ ), PCA outliers beyond 6SD of PC1 and PC2. Imputation: HRC.r1.1.2016 (predominantly European Ancestry), phasing using Eagle, Michigan Imputation Server.

## **VESTA cohort**

VESTA is a longitudinal behavior genetics study of cognitive and brain aging. The VETSA study design has several key characteristics. First, the sample has a narrow age range (~10 years), allowing for examination of individual differences in aging trajectories. Second, the initial assessment was in midlife (mean age 56; range 51-60), which provides a baseline for the transition to older age. Third, data previously collected on VETSA was obtained; of particular importance is a test of general cognitive ability administered at average age 20 and repeated in each wave of the study. Participants are members of the Vietnam Era Twin Registry, which is housed at the VA Puget Sound Health Care System in Seattle, WA, USA. All of the twins served the US military at some time during the Vietnam era (1965-1975). A 1992 study sought to recruit all Registry twins. It enrolled approximately 8000 individuals, including approximately 3300 twin pairs. VETSA participants were randomly recruited from those 3300 pairs. Eligibility for inclusion was based only on being 51-59 years old at the time of recruitment and willingness of both twins in a pair to participate. Both members of a pair did not need to participate to be included in wave 2 or wave 3. The average interval between assessment was approximately 6 years. Additional participants, including attrition replacement participants, were included at waves 2 and 3. Subsets have multi-modal MRI and neuroendocrine data. Data collection includes questionnaires filled out at home plus a daylong series of assessments. These include cognitive/neuropsychological assessment of multiple cognitive domains, personality and psychosocial assessments, and health/medical assessments. There are approximately 55% MZ and 45% DZ twins in the cohort. For cognitive, psychosocial, and health/medical data, there are 1291 individuals at wave 1, 1207 at wave 2 data, and 1196 at wave 3. Brain MRIs were obtained from 546 individuals at wave 1, 452 at wave 2, and 525 at wave 3. At wave 1 only, salivary cortisol, testosterone, and DHEAS data were collected on 780 participants.

VESTA participants live throughout the US. The sample is primarily Caucasian (European-American): 86% based on self-report. Only those of European-American ancestry based on genotype data were included in GWAS analyses. The average educational attainment is 13.8 years (SD=2.1). At wave 1, 79% were married, and 78% were employed full-time. Nearly 80% report no combat experience. The sample is similar with

respect to health and lifestyle characteristics to American men in their age range based on US Center for Disease Control and Prevention data<sup>13-15</sup>. This study was performed under appropriate oversight, including review boards at UC San Diego, Boston University, and VA Puget sound.

**NFL quantification: Nfl:** High throughput bioassays platforms or single analyte assays using the Quanterix Simoa HD-1 were used in this study. These human-specific immunoassays have been documented for measurements of these components in human plasma, as purified in Winston et al., 2016 and many other papers from the laboratory of Dr. Robert Rissman at UC San Diego. These assays are used routinely in the Rissman lab and all assays were performed according the manufacturer instructions.

**Genotyping and imputation:** The methods for genotyping, QC, and imputation have been described in detail elsewhere. See Logue et al. Mol Psych 2019 (PMCID: PMC6110977) for the details. In short, genotypes for VETSA participants were generated using the Illumina HumanOmniExpress-24 v1.0A BeadChip. Imputation was performed with 1000 genomes Phase 3 Reference data using MiniMac on the Michigan Imputation Server (<https://imputationserver.sph.umich.edu>). The analysis was done with RMW-RareMetalWorker v 4.13.7. Covariates included age, PCs 1-3, -- RMW already incorporates adjustments which take twinning into account including MZ/DZ relationships.

### **The Atherosclerosis Risk in Communities (ARIC) study**

The ARIC study, sponsored by the National Heart, Lung, and Blood Institute (NHLBI), is a prospective, bi-racial population-based study of atherosclerosis and cardiovascular diseases in 15,792 adults (11,478 non-Hispanic whites; 8,710 women) aged between 45 and 64 years at the baseline examination in 1987-1989. Participants were randomly selected and recruited from communities in Suburban Minneapolis, Minnesota; Washington County, Maryland; Forsyth County, North Carolina. In Jackson, Mississippi, only black residents were enrolled. Details about the ARIC study design and examination procedures have been previously published<sup>16</sup>. The ARIC study has been reviewed and approved by the Institutional Review Boards at all participating institutions. All participants provided written informed samples. This study uses data for participants who provided a blood draw at visit 3 (1993-1995) and who completed brain MRI and cognitive testing at visit 5 (2011-2013).

**NFL quantification:** NFL levels were quantified using the the Quanterix 4 plex and assays at University of Vermont by Russ Tracy.

**Genotyping:** Genotyping was performed using the Affymetrix GeneChip SNP Array 6.0 at Broad institute USA. Pre-imputation quality control include sample call rate > 0.95, and retaining SNP with call rate > 0.95, MAF > 0.01 and HWE p-value >  $10^{-5}$ . Imputation was performed using IMPUTE2 software on 1000G phase1 v3 reference panel.

## **CARDIA**

The CARDIA Study is a multisite population-based study conducted across four field centers: Birmingham, Alabama; Chicago, Illinois; Minneapolis, Minnesota; and Oakland, California. The study was established by the National Heart, Lung, and Blood Institute in 1984-1985. Black and White adults, between the ages of 18 to 30 years at the time of enrollment, were recruited. Participants have completed up to 9 examinations over 30 years. At the 25<sup>th</sup> year examination (2010-2011), participants had a blood draw and a subsample completed the brain MRI sub-study. This study uses data for this subset of participants.

The CARDIA study is a prospective, multi-center investigation of the natural history and etiology of cardiovascular disease in African Americans and whites 18-30 years of age at the time of initial examination. The initial examination included 5,115 participants selectively recruited to represent proportionate racial, gender, age, and education groups from four communities: Birmingham, AL; Chicago, IL; Minneapolis, MN; and Oakland, CA. Participants from the Birmingham, Chicago, and Minneapolis centers were recruited from the total community or from selected census tracts. Participants from the Oakland center were randomly recruited from the Kaiser-Permanente health plan membership. From the time of initiation of the study in 1985-1986, eight follow-up examinations have been conducted at years 2, 5, 7, 10, 15, 20, 25, and 30. Details of the study design have been published elsewhere<sup>17</sup>. The CARDIA study has been reviewed and approved by the Institutional Review Boards at

all participating institutions. At the 25<sup>th</sup> year examination (2010-2011), participants had a blood draw and a subsample completed the brain MRI sub-study. This study uses data for this subset of participants.

**NFL quantification:** NFL levels were quantified using the the Quanterix 4 plex and assays at University of Vermont by Russ Tracy.

**Genotyping:** Genotyping was performed using the Affymetrix 6.0 SNP array at Broad institute USA. Pre-imputation quality control include sample call rate > 0.95, and retaining SNP with call rate > 0.95, MAF > 0.01 and HWE p-value >  $10^{-5}$ . Imputation was performed using BEAGLE version 3.3.2 software on 1000G phase1 v3 reference panel.

**Supplementary Table 1: Demographic information for participating cohorts of European ancestry and African-American ancestry.**

	<b>Cohort</b>	<b>N</b>	<b>Age (mean, SD)</b>	<b>Female (%)</b>	<b>Plasma NfL (mean, SD)</b>	<b>Plasma NfL (median [25%-75%])</b>	<b>Log2 Plasma NfL (mean, SD)</b>	<b>Log2 Plasma NfL (median [25%-75%])</b>
<b>European</b>	<b>Rotterdam Study I</b>	2476	75.3 (6.11)	1445 (58.3)	17.74 (12.6)	14.99(11.32-20.23)	3.97(0.67)	3.91(3.50-4.34)
	<b>Rotterdam Study II</b>	1643	67.8 (7.12)	902 (55)	13.95 (13.66)	11.31(8.83-15.42)	3.59(0.68)	3.5(3.14-3.94)
	<b>Rhineland Study</b>	4019	55.4 (13.99)	2250 (56)	9.59 (8.64)	7.89 (5.72-11.47)	3.03 (0.76)	2.98 (2.56-3.52)
	<b>MEMENTO</b>	2195	70.8 (8.70)	1358 (62)	20.71 (12.95)	18.20 (13.40 – 25)	4.19 (0.72)	4.19 (3.74 – 4.64)
	<b>FHS cohort</b>	2048	70.6 (8.33)	1111(54.2)	22.90 (27.2)	17.8 (13.0-25.3)	4.24 (0.787)	4.15 (3.70-4.66)
	<b>BiDirect Study</b>	1899	52.1 (7.9)	921 (48.5)	9.77 (7.15)	8.45 (6.51-11.06)	2.16 (0.45)	2.13 (1.87-2.4)
	<b>CHS-EA</b>	1396	77.9 (4.3)	873 (63)	32.83 (36.0)	25.86 (20.2-34.49)	4.78 (0.72)	4.69 (4.34-5.11)
	<b>VESTA cohort</b>	828	67.5 (2.5)	0 (0)	13.5 (7.3)	11.9 (9.2- 15.3)	3.6 (0.61)	3.6 (3.2-3.9)
	<b>ARIC-EA</b>	823	63.1 (4.4)	477 (58)	19.4(45.9)	14.9(11.5-18.9)	3.95(0.71)	3.89( 3.52-4.24)
	<b>ADNI cohort</b>	578	85.3 (6.7)	243 (42)	42.8 (26.8)	36.9 (26.3 - 50.5)	5.2 (0.73)	5.2 (4.7-5.6)
	<b>CARDIA- EA</b>	340	51.0 (3.2)	179 (52.6)	9.88(5.87)	8.79(6.53-11.68)	3.14(0.65)	3.14(2.71-3.55)
	<b>ASPS-Fam</b>	287	64.3 (10.6)	172 (59.9)	35.27(16.42)	32.20 (22.80 - 44.10)	4.99(0.65)	5.01 (4.51-6.46)
<b>African American</b>	<b>ARIC-AA</b>	742	61.5 (4.5)	465 (62.7)	18.75( 47.95)	12.09( 9.24-17.35)	3.74(0.89)	3.60( 3.21-4.12)
	<b>CHS-AA</b>	273	76.3 (4.93)	177(65)	24.60 (16.23)	21.14 (16.18, 28.06)	4.43 (0.69)	4.4 (4.02-4.81)
	<b>CARDIA- AA</b>	127	48.9 (3.5)	72 (56.7)	10.24(7.56)	8.81(7.07-11.23)	3.17(0.66)	3.14(2.82-3.49)

**Abbreviations:** NfL, Neurofilament light; SD, Standard deviation; FHS, Framingham Heart Study; CHS, The Cardiovascular Health Study; ARIC, The Atherosclerosis Risk in Communities; VESTA, the Vietnam Era Twin Study of Aging; ADNI, the Alzheimer’s Disease Neuroimaging Initiative; CARDIA, the Coronary Artery Risk Development in Young Adults; ASPS-Fam, the Austrian stroke prevention family study; EA, European-American; AA, African-American

**Supplementary Table 2: Genome-wide significant loci (P-value < 5x10<sup>-8</sup>) associated with blood levels of Neurofilament light (NfL) in European and African-American Ancestry.**

SNV	Chromosome	Position*	Effect Allele	Other allele	Effect	SE	P-value	Nearest Genes	Genomic Locus	Direction	r <sup>2</sup>
<b>European ancestry</b>											
rs7203642	16	20367130	A	G	0.041	0.007	1.37x10 <sup>-8</sup>	UMOD	1	++++-+++++	15.9
rs12051560	17	70898201	A	G	0.033	0.006	9.94x10 <sup>-9</sup>	SLC39A11	2	++++-+++++	0
<b>African-American Ancestry</b>											
rs1026417	1	240437747	C	G	-0.433	0.076	1.36x10 <sup>-8</sup>	FMN2	1	---	0
rs17098087	12	63003902	C	G	0.440	0.074	2.59x10 <sup>-9</sup>	RP11-631N16.2	2	+++	0
rs73423978	12	88776247	T	C	0.332	0.060	3.50x10 <sup>-8</sup>	-	3	+++	53.9

\*base pair (BP) location is provided as per hg19

Note: The direction column represents the direction of  $\beta$  in individual cohorts: RS-I, RS-II, Rhineland study, CHS-EA, ASPS-Fam, VESTA, ADNI, ARIC-EA, FHS, BiDirect Study, CARDIA-EA, MEMENTO, and direction columns in African-American meta-analysis is the direction of  $\beta$  in following cohorts: ARIC-AA, CHS-AA, CARDIA-AA

**Supplementary Table 3: Genome-wide significant loci (P-value < 5x10<sup>-8</sup>) associated with blood levels of neurofilament light (NfL) in trans-ethnic GWAS meta-analysis**

SNV	Chromosome	Position	Effect Allele	Other Allele	P-value	Nearest gene	Direction	P-value Anc-Het	P-value Res-Het
rs11583796	1	60013252	C	T	1.14x10 <sup>-10</sup>	FGGY	--++++-++---	3.23x10 <sup>-12</sup>	0.215
rs67860565	4	86087697	G	A	4.72x10 <sup>-8</sup>	RN7SKP48	++++-++-----	1.93x10 <sup>-9</sup>	0.546
rs58152294	6	85482929	C	T	1.95x10 <sup>-8</sup>	TBX18	++-----+-----	1.05x10 <sup>-9</sup>	0.086

\*base pair (BP) location is provided as per hg19

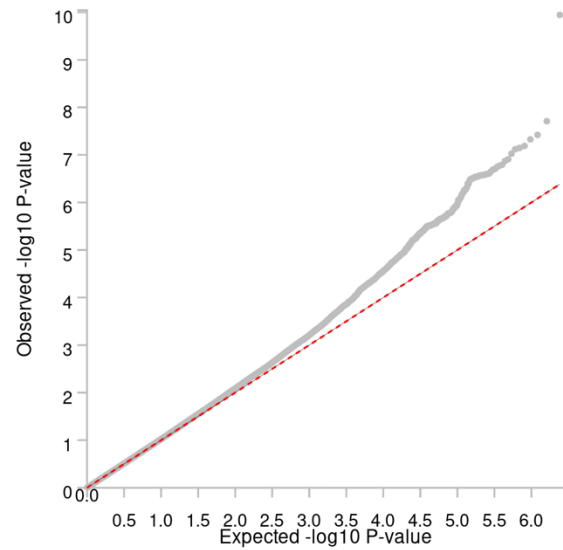
Note: The direction column represents the direction of  $\beta$  in individual cohorts: RS-I, RS-II, Rhineland study, CHS-EA, ASPS-Fam, VESTA, ADNI, ARIC-EA, FHS, BiDirect Study, CARDIA-EA, MEMENTO, ARIC-AA, CHS-AA, CARDIA-AA

Abbreviations: SNV, single nucleotide variation, Anc-Het, Ancestral Heterogeneity; Res-Het, Residual Heterogeneity.

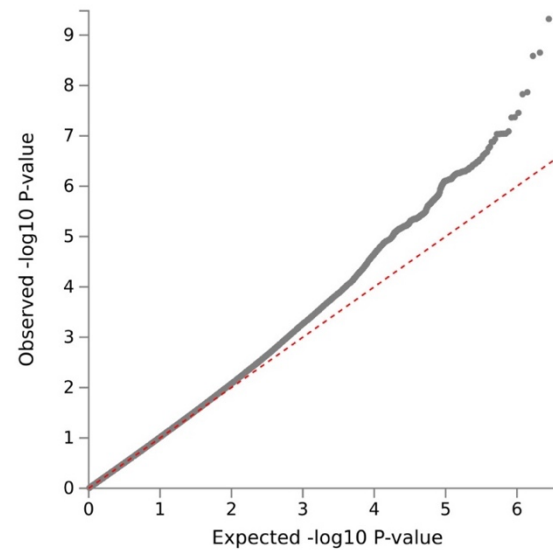


## Supplementary Figures:

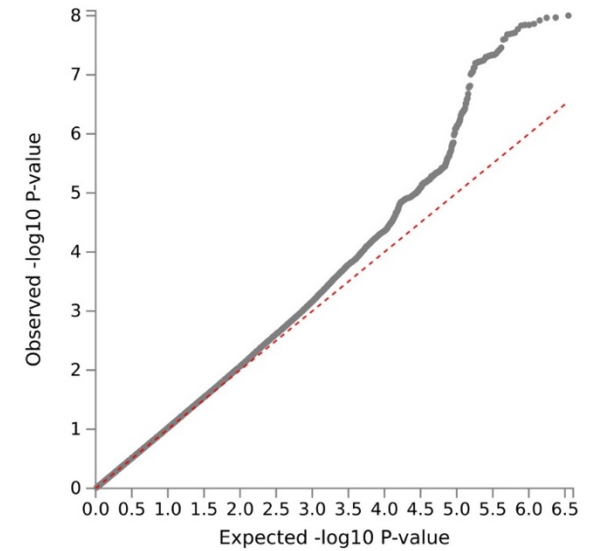
(A)



(B)

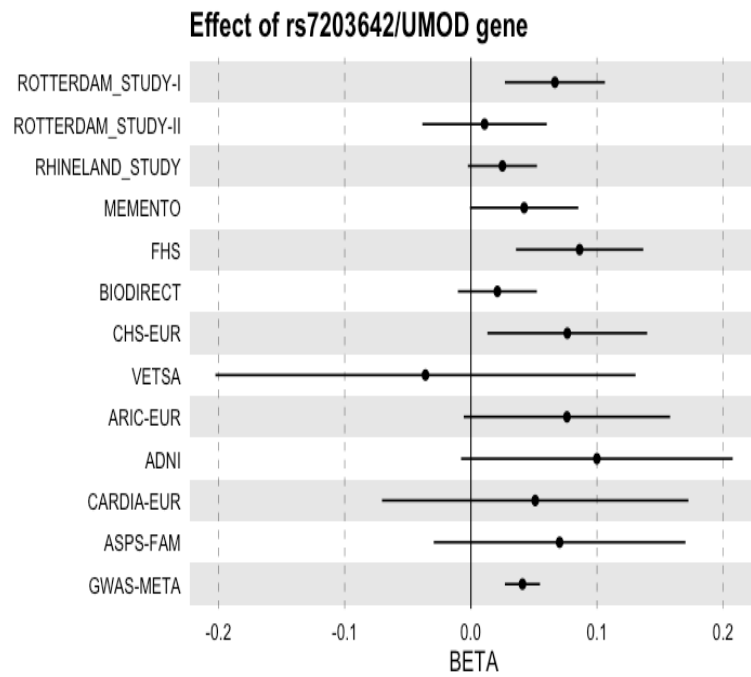


(C)

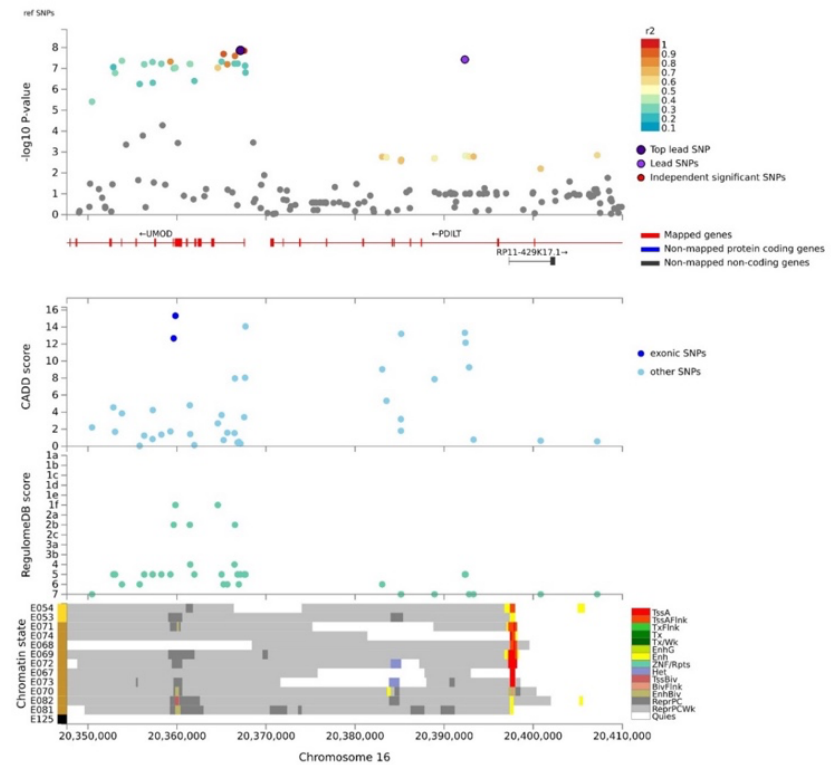


**Supplementary Figure 1:** Quantile-Quantile (Q-Q) plot for GWAS meta-analysis for European ancestry (A) African-American Ancestry (B) and Trans-ethnic meta-analysis (C). The grey dotted line represents the distribution of observed  $-\log_{10}(p\text{-values})$  against the theoretical model distribution of expected  $-\log_{10}(p\text{-values})$ . The red dotted line represents the theoretical model distribution of expected  $-\log_{10}(p\text{-values})$  under the null distribution.

A)

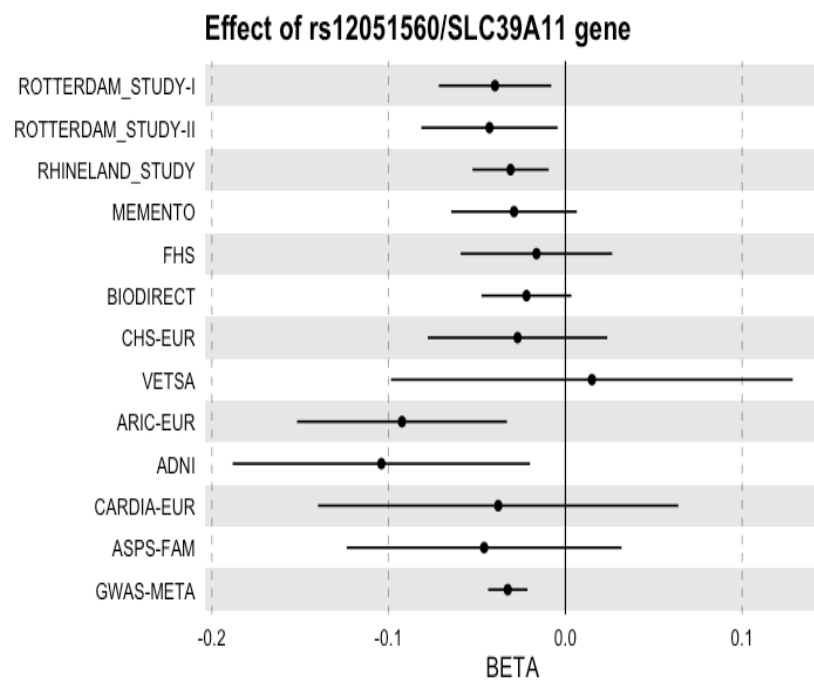


B)

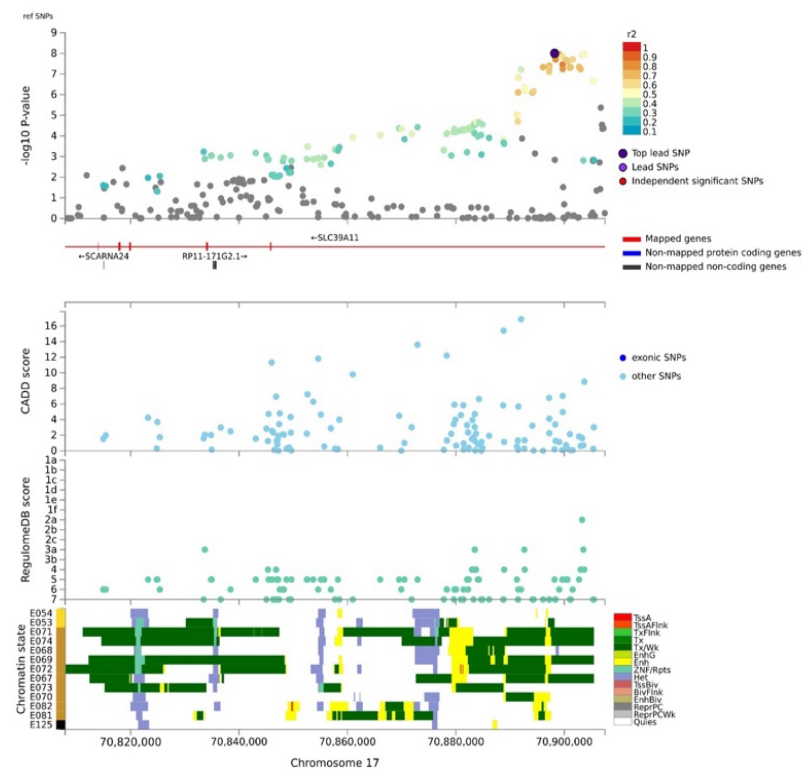


**Supplementary Figure 2: (A)** Forest plot of the lead SNP in *UMOD* locus **(B)** Regional plot showing the top lead genetic variant inside *UMOD* gene, their GWAS *P-value* as  $-\log_{10}$ , Combined Annotation Dependent Depletion (CADD) score, Regulome Data Base score and expression Quantitative loci (eQTLs). The regional plot with annotation information is generated using the Functional Mapping and Annotation of Genome-Wide Association Studies (FUMA) platform.

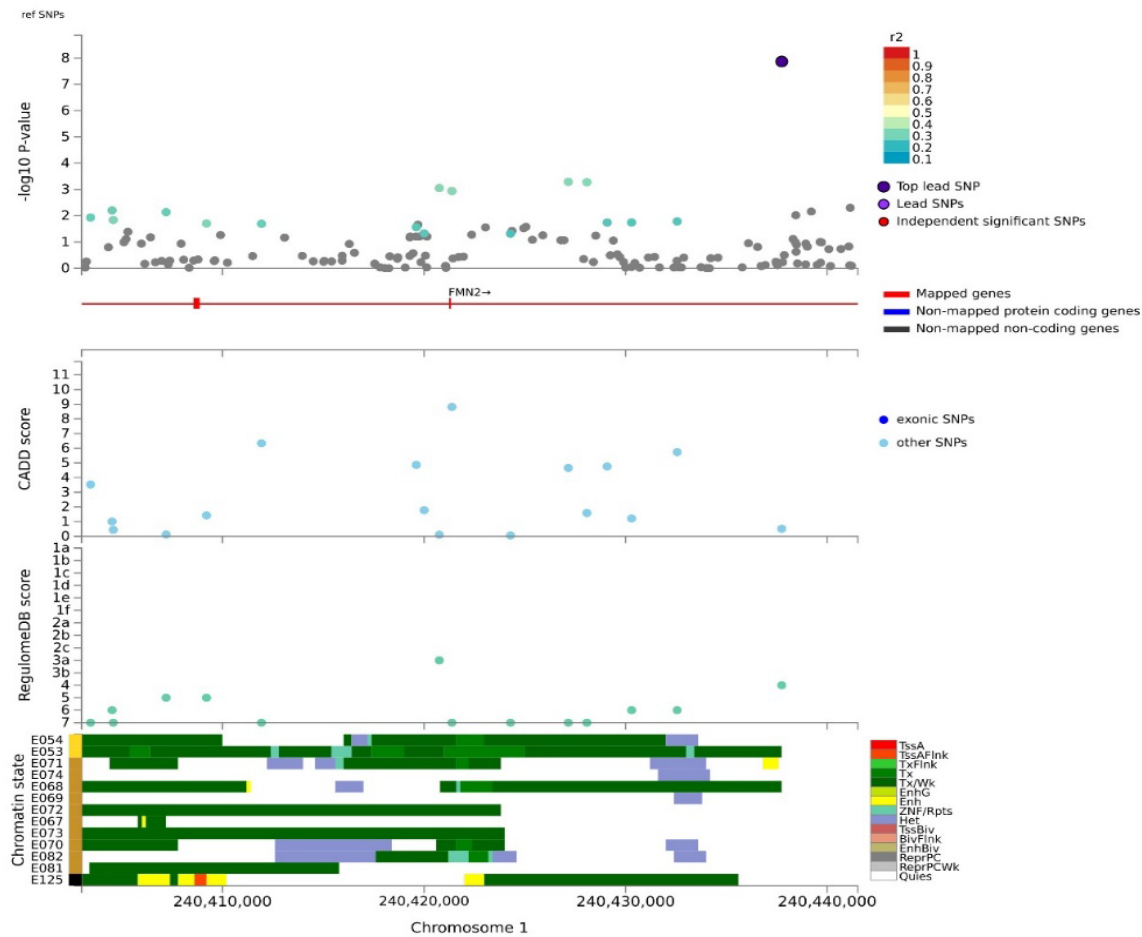
A)



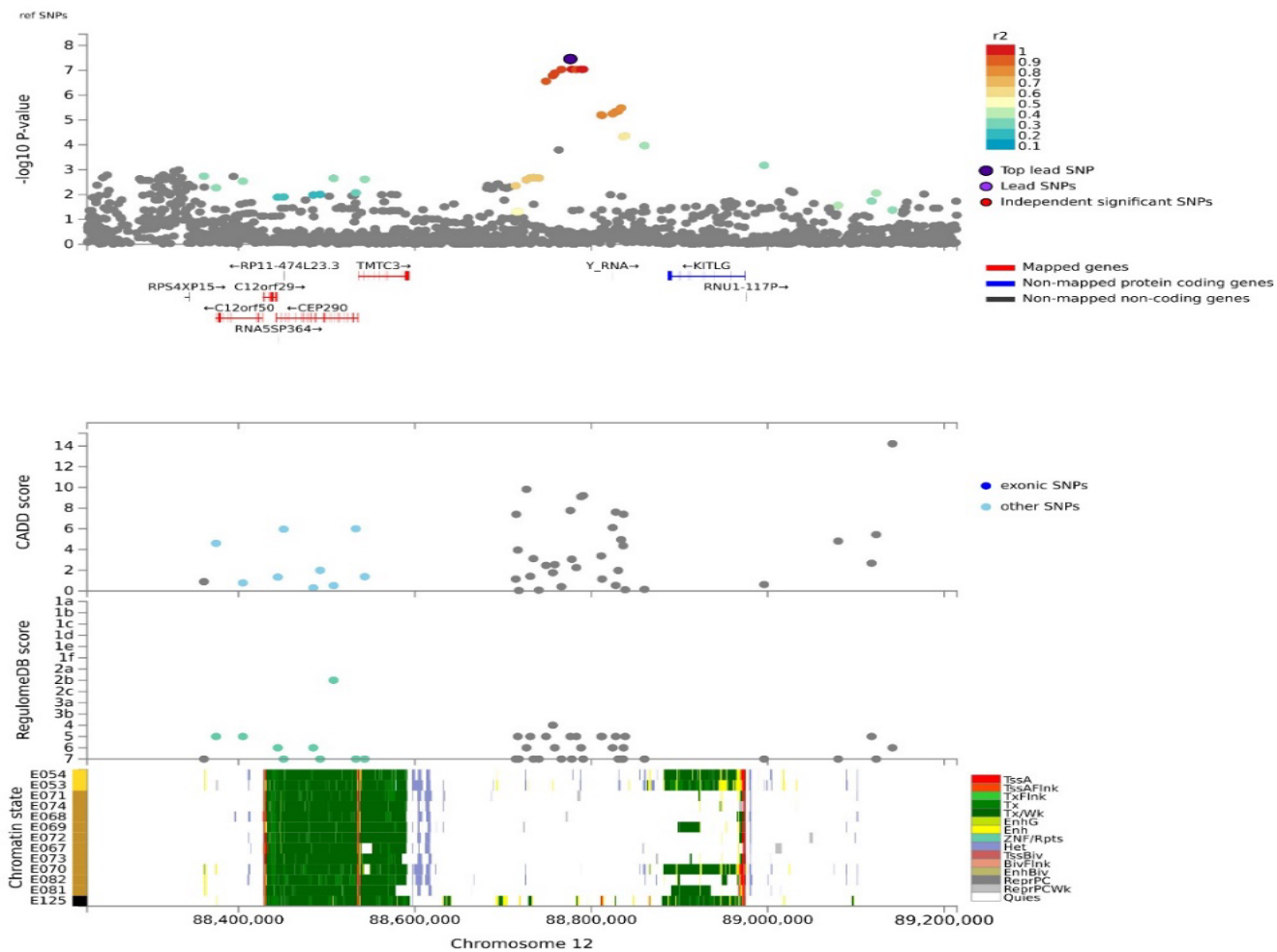
B)



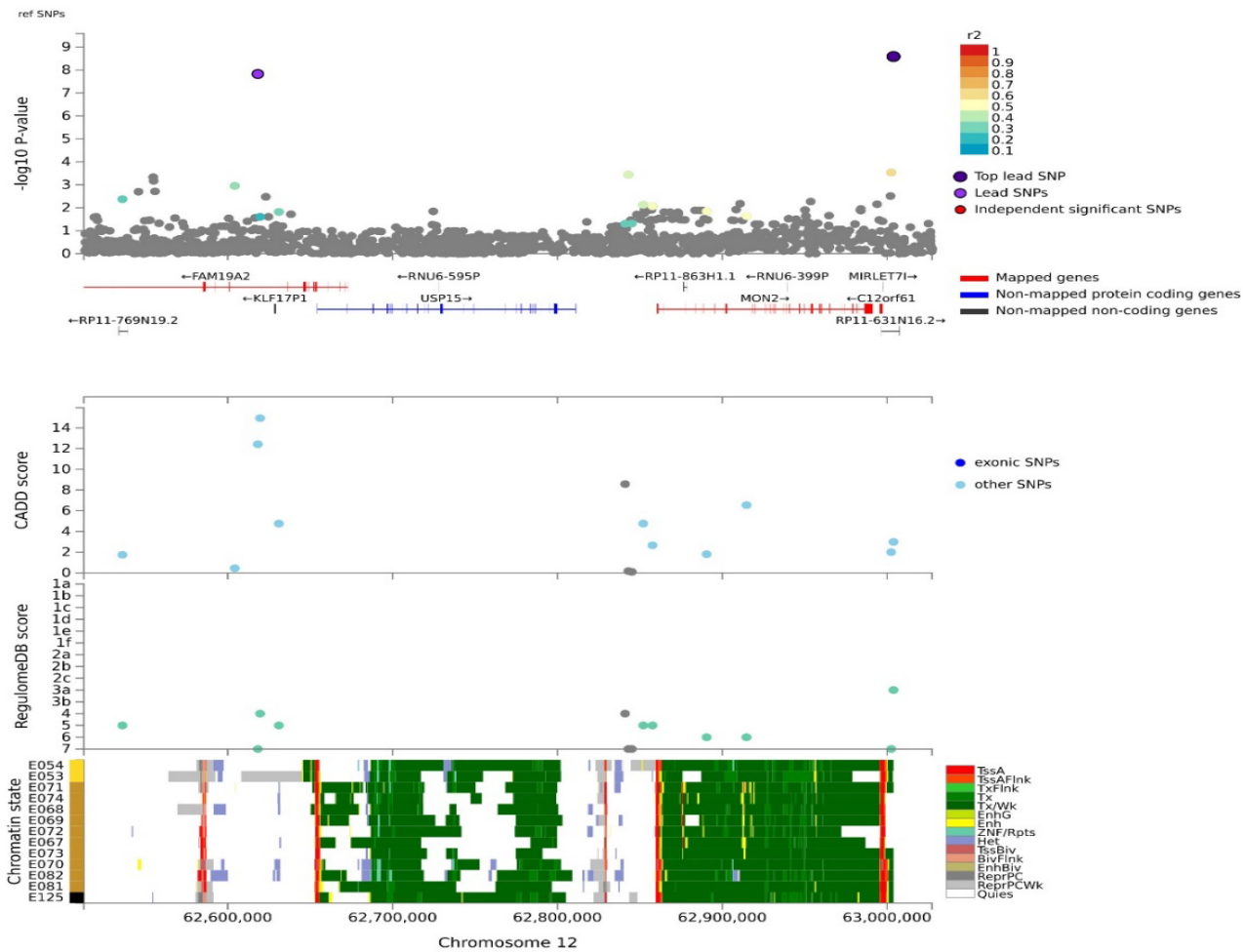
**Supplementary Figure 3: (A)** Forest plot of the lead SNP in *SLC39A11* locus **(B)**Regional plot showing the top Lead SNP inside *SLC39A11* gene, their GWAS P-value as  $-\log_{10}$ , Combined Annotation Dependent Depletion (CADD) score, Regulome Data Base score and expression Quantitative loci (eQTLs). The regional plot with annotation information is generated using the Functional Mapping and Annotation of Genome-Wide Association Studies (FUMA) platform.



**Supplementary Figure 4:** Regional plot for locus 1q43 in the African-American ancestry, their GWAS P-value as  $-\log_{10}$ , Combined Annotation Dependent Depletion (CADD) score, Regulome Data Base score and expression Quantitative loci (eQTLs). The regional plot with annotation information is generated using the Functional Mapping and Annotation of Genome-Wide Association Studies (FUMA) platform.

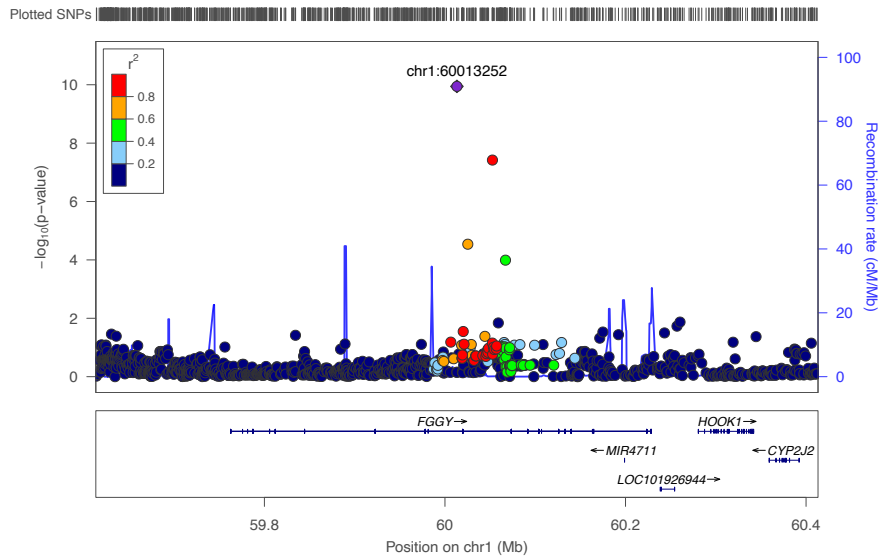


**Supplementary Figure 5:** Regional plot for locus 12q14 in the African-American ancestry, their GWAS P-value as  $-\log_{10}$ , Combined Annotation Dependent Depletion (CADD) score, Regulome Data Base score and expression Quantitative loci (eQTLs). The regional plot with annotation information is generated using the Functional Mapping and Annotation of Genome-Wide Association Studies (FUMA) platform.

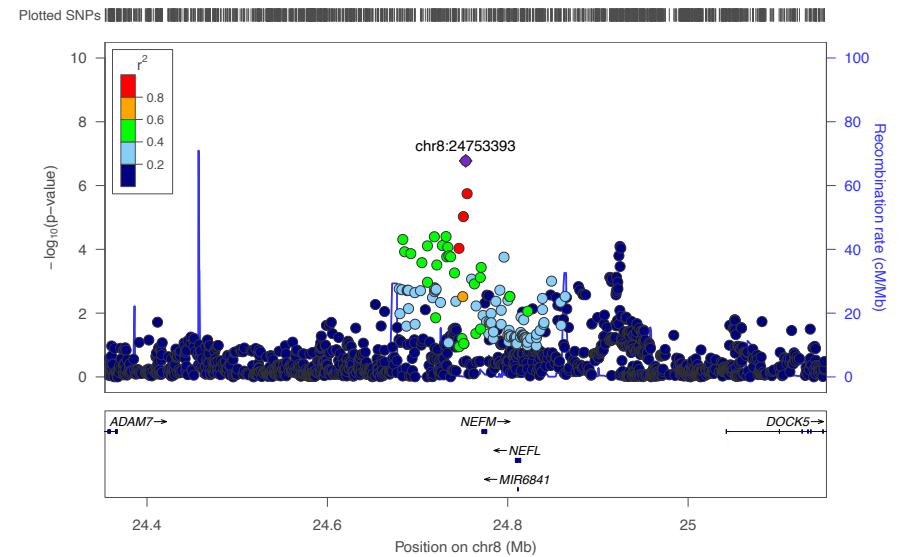


**Supplementary Figure 6:** Regional plot for locus 12q21 in the African-American ancestry, their GWAS P-value as  $-\log_{10}$ , Combined Annotation Dependent Depletion (CADD) score, Regulome Data Base score and expression Quantitative loci (eQTLs). The regional plot with annotation information is generated using the Functional Mapping and Annotation of Genome-Wide Association Studies (FUMA) platform.

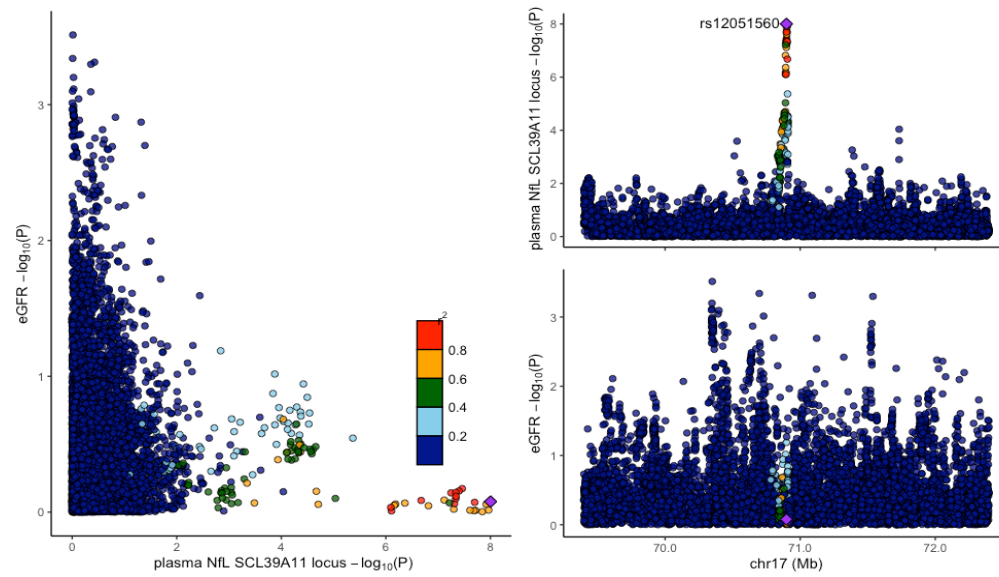
(A)



(B)



**Supplementary Figure 7:** Regional plot for two loci in *FGGY* (A) and locus near *NEFM* gene (B) identified in the meta-analysis of neurofilament light (NfL) genome-wide association study (GWAS) in European ancestry. The genetic variants are denoted as colored circles with their P-values ( $-\log_{10}$ ) on the left Y-axis and genomic location is based on build 37 on X-axis. Lead SNPs (purple diamond) are marked with their genomic location. Recombination rates are plotted on right Y-axis to represent the local linkage disequilibrium (LD) structure. The LD between the genetic variants is provided with a color scale, ranging from blue ( $r^2 = 0$ ) to red ( $r^2 = 1$ ). LD calculations are based on 1000 genome European Ancestry.



**Supplementary Figure 8:** Colocalization overview of *SLC39A11* locus with kidney function in European ancestry.



## Supplementary References

1. Rissin, D.M. *et al.* Simultaneous detection of single molecules and singulated ensembles of molecules enables immunoassays with broad dynamic range. *Anal Chem* **83**, 2279-85 (2011).
2. Rohrer, J.D. *et al.* Serum neurofilament light chain protein is a measure of disease intensity in frontotemporal dementia. *Neurology* **87**, 1329-36 (2016).
3. Fani, L., Ahmad, S., Ikram, M.K., Ghanbari, M. & Ikram, M.A. Immunity and amyloid beta, total tau and neurofilament light chain: Findings from a community-based cohort study. *Alzheimer's & Dementia* **17**, 446-456 (2021).
4. Teismann, H. *et al.* Establishing the bidirectional relationship between depression and subclinical arteriosclerosis--rationale, design, and characteristics of the BiDirect Study. *BMC Psychiatry* **14**, 174 (2014).
5. Schmidt, R., Fazekas, F., Kapeller, P., Schmidt, H. & Hartung, H.P. MRI white matter hyperintensities: three-year follow-up of the Austrian Stroke Prevention Study. *Neurology* **53**, 132-9 (1999).
6. Schmidt, R. *et al.* Assessment of cerebrovascular risk profiles in healthy persons: definition of research goals and the Austrian Stroke Prevention Study (ASPS). *Neuroepidemiology* **13**, 308-13 (1994).
7. Ghadery, C. *et al.* R2\* mapping for brain iron: associations with cognition in normal aging. *Neurobiol Aging* **36**, 925-32 (2015).
8. Pirpamer, L. *et al.* Determinants of iron accumulation in the normal aging brain. *Neurobiol Aging* **43**, 149-55 (2016).
9. Norgren, N., Karlsson, J.E., Rosengren, L. & Stigbrand, T. Monoclonal antibodies selective for low molecular weight neurofilaments. *Hybrid Hybridomics* **21**, 53-9 (2002).
10. Disanto, G. *et al.* Serum Neurofilament light: A biomarker of neuronal damage in multiple sclerosis. *Ann Neurol* **81**, 857-870 (2017).
11. Khalil, M. *et al.* Serum neurofilament light levels in normal aging and their association with morphologic brain changes. *Nat Commun* **11**, 812 (2020).
12. Dufouil, C. *et al.* Cognitive and imaging markers in non-demented subjects attending a memory clinic: study design and baseline findings of the MEMENTO cohort. *Alzheimers Res Ther* **9**, 67 (2017).
13. Kremen, W.S. *et al.* Genetic and environmental influences on the size of specific brain regions in midlife: the VETSA MRI study. *Neuroimage* **49**, 1213-1223 (2010).
14. Kremen, W.S., Franz, C.E. & Lyons, M.J. VETSA: the Vietnam era twin study of aging. *Twin Research and Human Genetics* **16**, 399-402 (2013).
15. Kremen, W.S., Franz, C.E. & Lyons, M.J. Current status of the Vietnam era twin study of aging (VETSA). *Twin Research and Human Genetics* **22**, 783-787 (2019).
16. Wright, J.D. *et al.* The ARIC (atherosclerosis risk in communities) study: JACC focus seminar 3/8. *Journal of the American College of Cardiology* **77**, 2939-2959 (2021).
17. Friedman, G.D. *et al.* CARDIA: study design, recruitment, and some characteristics of the examined subjects. *Journal of clinical epidemiology* **41**, 1105-1116 (1988).